

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

DATA ARCHIVE SYSTEM

Patent Number: JP10254629
Publication date: 1998-09-25
Inventor(s): OEDA TAKASHI; TOMONO YOJI; TSUJIOKA SHIGEO; YAMAMOTO AKIRA
Applicant(s):: HITACHI LTD
Requested Patent: ☐ JP10254629
Application Number: JP19970060521 19970314
Priority Number(s):
IPC Classification: G06F3/06 ; G06F3/06 ; G06F12/16
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To make it possible to take out and carry archived data without transferring the archived data through a network by allowing a library controller to provide a control means which carries a storage medium to a medium discharge/insertion slot for a changer at a request from a client computer.

SOLUTION: A portable medium library device 200 has the medium discharge/insertion slot 210 and discharges a storage medium having data that a client computer specifies. The discharged storage medium is read and written by a drive device of the client computer or a drive device of a remote computer. The portable medium storage library 200 consists of a library controller 220, a drive device 280, a medium insertion slot 210, a changer 290, and a medium storage box 295. In this case, the changer 290 carries the storage medium stored in the medium storage box 295 to drive devices 280-1 to 280-5 or the medium insertion/discharge slot 210 according to an instruction that the library controller 290 sends out.

Data supplied from the esp@cenet database - I2

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-254629

(43) 公開日 平成10年(1998) 9月25日

(51) Int.Cl. ⁶	識別記号	F I
G 0 6 F 3/06	3 0 1	G 0 6 F 3/06
	5 4 0	3 0 1 Z
12/16	3 2 0	5 4 0
		3 2 0 L

審査請求 未請求 請求項の数 7 O L (全 16 頁)

(21) 出願番号 特願平9-60521

(22) 出願日 平成9年(1997) 3月14日

(71) 出願人 000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地

(72) 発明者 大枝 高
神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

(72) 発明者 伴野 洋二
東京都品川区南大井6丁目27番18号 株式会社日立製作所情報事業本部内

(72) 発明者 辻岡 重夫
神奈川県海老名市下今泉810番地 株式会社日立製作所オフィスシステム事業部内

(74) 代理人 弁理士 小川 勝男

最終頁に続く

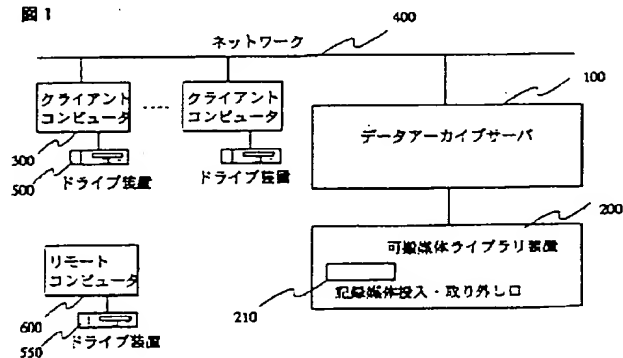
(54) 【発明の名称】 データアーカイブシステム

(57) 【要約】

【課題】 集中型のデータアーカイブサーバはネットワークでデータを転送する必要があるため大容量のデータを利用する場合、低速になってしまう。特に、比較的通信速度が遅いか通信手段がない遠隔地から利用する場合に困難であった。

【解決手段】 可搬媒体ライブラリ装置で、クライアントの要求する媒体を取り出して使用可能とする。クライアントはローカルなドライブ装置でその媒体を利用する。ライブラリ装置では記憶媒体を冗長性を付加し、排出された媒体に対するアクセスを処理する。

図 1



【特許請求の範囲】

【請求項 1】クライアントコンピュータとデータアーカイブサーバと可搬媒体ライブラリ装置で構成されるデータアーカイブシステムにおいて、

前記クライアントコンピュータは前記可搬媒体ライブラリ装置の記憶媒体をリードできるドライブ装置を備え、少なくとも一台のクライアントコンピュータがネットワークで前記データアーカイブサーバに接続され、

前記データアーカイブサーバは前記可搬媒体ライブラリ装置と接続され、

前記可搬媒体ライブラリ装置は、前記記憶媒体と前記記憶媒体を収納する媒体収納庫と前記記憶媒体を装填してデータの読みだしや書き込みを行うドライブ装置と前記記憶媒体を前記可搬媒体ライブラリ装置へ投入もしくは前記可搬媒体ライブラリ装置から排出するための媒体投入・排出口と前記記憶媒体を搬送する媒体チェンジャを備え、

前記ネットワーク経由で前記クライアントコンピュータから前記データアーカイブサーバへ送られ、前記可搬媒体ライブラリ装置の記憶媒体に蓄積されたデータを、当該記憶媒体を前記可搬媒体ライブラリ装置の前記媒体投入・排出口から排出し、前記クライアントコンピュータの前記ドライブ装置に装填することにより利用可能とすることを特徴とするデータアーカイブシステム。

【請求項 2】請求項 1 に示すデータアーカイブシステムで、

前記記憶媒体は複数枚の組でバリティグループを構成し、

前記バリティグループのなかの任意の記憶媒体が障害をおこしてもデータを回復することができる冗長性を持ち、

前記可搬媒体ライブラリ装置は前記冗長性を用いて障害を起こした記憶媒体のデータを回復する手段を持ち、

前記クライアントコンピュータからデータアーカイブサーバ経由で排出された記憶媒体のデータへの読み書きを可能とすることを特徴とするデータアーカイブシステム。

【請求項 3】クライアントコンピュータとデータアーカイブサーバと可搬媒体ライブラリ装置で構成されるデータアーカイブシステムにおいて、

前記クライアントコンピュータは前記可搬媒体ライブラリ装置の記憶媒体をリードできるドライブ装置を備え、少なくとも一台のクライアントコンピュータがネットワークで前記データアーカイブサーバに接続され、

前記データアーカイブサーバは前記可搬媒体ライブラリ装置と接続され、

前記可搬媒体ライブラリ装置は、前記記憶媒体と前記記憶媒体を収納する媒体収納庫と前記記憶媒体を装填してデータの読みだしや書き込みを行うドライブ装置と前記記憶媒体を前記可搬媒体ライブラリ装置へ投入もしくは

前記可搬媒体ライブラリ装置から排出するための媒体投入・排出口と前記記憶媒体を搬送する媒体チェンジャを備え、

前記ネットワーク経由で前記クライアントコンピュータから前記データアーカイブサーバへ送られ、前記可搬媒体ライブラリ装置の記憶媒体に蓄積されたデータを、当該記憶媒体のデータを別の記憶媒体にコピーし、コピーした記憶媒体を前記可搬媒体ライブラリ装置の前記媒体投入・排出口から排出し、前記クライアントコンピュータの前記ドライブ装置に装填することにより利用可能とすることを特徴とするデータアーカイブシステム。

【請求項 4】請求項 2 に記載のデータアーカイブシステムで、

排出中の前記記憶媒体へのデータ更新情報を保持することにより、記憶媒体の再投入後に排出中のデータ更新を当該記憶媒体上に反映することのできるデータアーカイブシステム。

【請求項 5】請求項 2 に記載のデータアーカイブシステムで、

排出した前記記憶媒体を装填した前記クライアントコンピュータとデータアーカイブサーバはネットワークで接続されていないことを特徴とするデータアーカイブシステム。

【請求項 6】請求項 2 に記載のデータアーカイブシステムで、

排出した前記記憶媒体を装填した前記クライアントコンピュータとデータアーカイブサーバは通信手段をもち、前記データアーカイブサーバもしくは前記可搬媒体ライブラリ装置は排出中の記録媒体へのデータ更新情報を保持し、

前記クライアントコンピュータが前記排出された記憶媒体へアクセスする場合、前記通信手段を用いアクセスするデータへの更新の有無を確認し、更新されていた場合更新されたデータを前記通信手段を使って取得するか、または、更新されていることを表示することを特徴とするデータアーカイブシステム。

【請求項 7】請求項 6 に記載のデータアーカイブシステムで、

排出した前記記憶媒体を装填した前記クライアントコンピュータとデータアーカイブサーバは通信手段をもち、前記クライアントコンピュータは前記可搬媒体ライブラリ装置から排出された前記記録媒体へのデータ更新情報を保持し、

前記データアーカイブサーバが前記排出された記憶媒体へアクセスする場合、前記通信手段を用いアクセスするデータへの更新の有無を確認し、更新されていた場合更新されたデータを前記通信手段を使って取得するか、または、更新されていることを報告することを特徴とするデータアーカイブシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、可搬記憶媒体を対象としたライブラリ型記憶装置を利用したデータアーカイブシステムに関する。特に、可用性向上のために、ライブラリ型記憶装置に冗長性をもたせたデータアーカイブシステムに関する。

【0002】

【従来の技術】従来、コンピュータシステムのデータアーカイブ方法には、各コンピュータ毎にアーカイブ装置を持ち個別に行う方法と、データアーカイブサーバを持ち集中的に行う方法があった。

【0003】前者の分散型で処理する方法は、各コンピュータに磁気テープ装置などの記憶装置を接続し、当該コンピュータ上のデータを接続された磁気テープ装置などに格納する方法である。この方法は格納したデータを持ち運ぶ際、必要なデータを格納した磁気テープを持っていけばよい。データを持ち運ぶための処理時間が短くてすむという利点があるが、各コンピュータ毎にアーカイブ用のデータ記憶装置を設置する必要があるためコストが高いという欠点があった。また、データを格納した記憶媒体の管理も自動化されていないため、データを紛失しやすく、必要なデータの取り出し・検索に時間がかかるという問題もあった。さらに、複数のユーザでデータを共有するのが困難であるという問題もあった。

【0004】後者のデータアーカイブサーバで集中的に処理する方法は、複数の（クライアント）コンピュータとデータアーカイブサーバをネットワークで接続し、データアーカイブサーバが管理する記憶装置に集中的にデータを格納する方法である。この方法はクライアント一台あたりのデータアーカイブに要するコストを下げることができる。そこで、特公平7-27442”データ記憶装置階層構造におけるヒット率を向上させる方法およびそのための装置”に示すような比較的単価の高いライブラリ型記憶装置なども利用することができるようになる。このようなライブラリ型記憶装置を使用すれば、記憶媒体の管理を自動化でき、データの紛失や、データ取り出し・検索に時間がかかるという前者の方法が持っている問題点を解決することができる。つまり、クライアントコンピュータでは大容量すぎてオフラインにしか持てなかったデータを、ネットワークを介してではあるがオンライン的に持つことができるようになる。さらに、複数のクライアントコンピュータから容易にライブラリ型記憶装置に格納されているデータを共用することができる。光磁気ディスクやDVD（Digital Video Disk）など、ランダムアクセスが可能であり、かつ磁気ディスクなどに比べビットコストの安い可搬媒体記憶装置が出現してきたこともあり、今後ライブラリ型記憶装置を利用したデータアーカイブシステムが普及することが予想される。

【0005】次に、データの可用性を高めるために冗長

性を持たせた記憶装置に関する従来技術について説明する。

【0006】以下に示すPattersonの論文が知られている。

【0007】エー．シー．エム．シグモッド コンファレンス プロシーディング、1988年、6月、ページ109-116 (D.Patterson, et al: A Case for Redundant Arrays of Inexpensive Disks(RAID), ACM SIGMOD conference proceeding, Chicago, IL, June 1-3, 1988, pp.109-116)

Pattersonの論文は、ディスクアレイ上のデータ配置に関する技術を開示したものである。

【0008】ディスクアレイは、ディスクシステムの高性能化、高信頼化を実現するための機構である。ディスクアレイでは、高性能化のために、物理的には複数のディスク装置を、処理装置に対しては1台のディスク装置に見せかける。一方、高信頼化のためには、データを格納したディスク装置に障害が発生した場合、データの回復を行うための冗長データを別のディスク装置に格納しておく。

【0009】通常、ディスク装置のリード／ライト単位となるデータをレコードと呼ぶが、Pattersonの論文では、いくつかのレコード配置方法が提案されている。ただし、ディスクアレイを用いた場合、処理装置から見たリードライト単位であるレコードと、ディスク装置に実際に記録されるレコードとはデータ長が異なる場合がある。以下、前者を論理レコード、後者を物理レコードと呼ぶ。以下、Pattersonの論文で提案されているいくつかのレコード配置方法の説明を行う。

【0010】第1の配置方法は、論理レコード、すなわち、処理装置側から見たレコードを、ディスク装置上では、 m 個 ($m \geq 1$) の物理レコードに分割して格納する配置方法である。以下、この配置方法を、分割配置方法と呼ぶ。分割配置を用いた場合、1つの論理レコードを m 台のディスク装置との間で転送できることから、見かけ上データ転送速度を m 倍に向上させたのと同様の効果を得ることができる。

【0011】次に、分割配置における冗長データの作成方法を説明する。分割配置では、論理レコードを分割した m 個の物理レコードに対し、 n 個 ($n \geq 1$) の冗長データが作成され、それぞれを、1つの物理レコード（全体で n 個ある）としてディスク装置に格納する。以下、処理装置が直接リード／ライトするデータを格納した物理レコードをデータレコード、冗長データを格納した物理レコードをパリティレコードと呼ぶ。また、 m 個のデータレコードと n 個のパリティレコードから構成されるグループを、パリティグループと呼ぶ。通常、パリティグループ内のパリティレコードの数が n 個であれば、 n 台までのディスク装置に障害が発生してもそのパリティ

グループのデータは回復可能である。

【0012】第2の配置方法は、処理装置から見たリード／ライト単位である論理レコードを、1つの物理レコード、すなわち、1つのデータレコードとして、ディスク装置上に格納する配置方法である。以下、これを非分割配置と呼ぶ。したがって、論理レコードは、データレコードと等価なる。(それぞれの物理レコードには、データレコードあるいはパリティレコードが割り当てられるため、物理レコードと論理レコードは必ずしも等価にならない。すなわち、1つの論理レコードは、1つの物理レコードではあるが、1つの物理レコードは、1つの論理レコードであるというわけではないし、パリティレコードである場合もある。)非分割配置の特長は、ディスクアレイを構成するそれぞれのディスク装置ごとにリード／ライト処理が実行可能な点である。(分割配置方法をとると、リード／ライトのために複数のディスク装置を専有する必要がある。)したがって、非分割配置をとると、ディスクアレイ内で実行できるリード／ライト処理の多重度を向上させることが可能となり、性能向上を実現できる。非分割配置でも、m個のデータレコードから、n個のパリティレコードを作成し、ディスク装置に格納される。ただし、分割配置の場合、パリティグループ内のデータレコードの集合が、処理装置から見た1つの論理レコードを形成するのに対し、非分割配置の場合、データレコードのそれぞれが、処理装置から見るとまったく独立した論理レコードとなる。

【0013】計算機システムにおいて、ディスク装置外にしばしば用いられる記憶装置として、磁気テープ、光記憶装置等がある。特に最近では、DVD(Digital Video Disk)が注目されている。これらの記憶装置の特徴は、いずれも記憶媒体とドライブ装置が分離されており、記憶媒体を任意のドライブ装置に装填し、記憶媒体上のデータを読み書きするという点である。一般にこれらの媒体は、可搬媒体と呼ばれる。大規模コンピュータシステムにおいては、非常に多くの枚数の可搬媒体の管理を容易に実現するために、ライブラリが導入される。ライブラリには、記憶媒体、ドライブ装置(以下ドライブと表記)に加えて、多くの枚数の記憶媒体を収納する収納庫と、収納庫とドライブ装置との間で、記憶媒体を転送する媒体チェンジャ(以下チェンジャと表記)などが含まれる。

【0014】計算機システムで扱うデータは、ますます大規模化しているため、その可用性の向上に対するニーズも非常に高い。したがって、上記のような可搬媒体で構成される記憶装置システムにおいても、Pattersonの論文で提案されているようなコンセプトを適用することにより、高可用性を実現することは有効である。

【0015】

【発明が解決しようとする課題】自動記憶装置ライブラ

リを利用したデータアーカイブサーバは、上記のような利点があるが、一方格納したデータを利用しようとするとネットワーク経由でデータをコピーする必要がある。一般にネットワークを介したデータ転送は低速であり、大容量なデータの転送は時間がかかる。例えば、現在普及しているイーサネット(10Mビット毎秒)を用いて、パーソナルコンピュータで一般に用いられている磁気ディスク装置(1Gバイト)にデータを転送する場合、少なくとも800秒以上の時間を要する。実際にはネットワークの効率、ネットワーク以外の処理時間にかかるためこの数倍の時間がかかることになる。特に、遠隔地からデータを利用する場合はネットワークがつながっていても、通常はローカルなネットワークに比べずっと低速であるし(公衆回線を用いる標準的なモデムは28.8kビット毎秒)、また、ネットワークが利用できない場合も多い。そこで、可換媒体(磁気テープ、光ディスクなど)でデータを持ち運ぶことになる。

【0016】ところが、ライブラリ型記憶装置は通常媒体を取り出して他の装置で利用するための媒体排出要求をクライアントコンピュータから受け付ける手段がなく、必要な媒体を取り出して持っていくことができない。

【0017】さらに、もし媒体を取り出した場合、その媒体上のデータはアクセスできなくなってしまう、各クライアントコンピュータが共通的にアクセスするデータが格納されている媒体が取り外された場合、そのデータの読み出し、更新、追加などの処理はその媒体が元通りにセットされるまでできなくなってしまう。

【0018】本発明の第一の目的は、アーカイブデータのネットワーク転送処理なしでアーカイブしたデータを取り出し、持ち運べる集中型データアーカイブサーバシステムを提供することにある。

【0019】本発明の第二の目的は、ライブラリ装置を利用したデータアーカイブサーバで記憶媒体が取り外されても当該媒体上のデータへのアクセスを処理することのできるシステムを提供することにある。

【0020】本発明の第三の目的は、上記のシステムで取り出した記憶媒体のデータに対する更新と取り出し中のデータアーカイブサーバ側でのデータ更新を、その記憶媒体がライブラリ記憶装置に再投入された後に同期化する手段を提供することにある。

【0021】本発明の第四の目的は、上記のシステムで取り出した記憶媒体のデータに対する更新と取り出し中のデータアーカイブサーバ側でのデータ更新を、リアルタイムで同期化する手段を提供することにある。

【0022】

【課題を解決するための手段】可搬媒体自動ライブラリ記憶装置を利用した集中型アーカイブサーバで、媒体を抜きだし携帯できるようにする。原本の媒体を持ち出す場合とコピーを作成して持ち出す場合あり。

【0023】上記第一の目的のために、本発明では可搬媒体ライブラリ装置に媒体排出・投入口を設け、クライアントコンピュータからの要求によりライブラリ制御装置はチェンジャに記憶媒体を媒体排出・投入口に搬送する制御手段を設けた。クライアントが要求するデータが格納されている記憶媒体そのものを排出する方法と排出用の記憶媒体を可搬媒体ライブラリ装置に備え、ライブラリ制御装置が記憶媒体の複製を制御する手段を持ち、記憶媒体の複製を作成して排出する方法がある。また、クライアントコンピュータには記憶媒体のデータを読むことのできるドライブ装置もしくは読み書きができるドライブ装置を備えることとした。

【0024】上記第二の目的のために、ライブラリ制御装置に記憶媒体の冗長性を付加した組（バリティグループ）を管理する構成情報管理手段と、付加した冗長性を用いて障害を起こした記憶媒体があってもクライアントコンピュータからのアクセスを可能にする排出・複製媒体アクセス制御手段を設け、排出された記憶媒体へのアクセスを可能とした。

【0025】上記第三の目的のために、ライブラリ制御装置に排出された記憶媒体への更新情報を保持する手段と、排出された記憶媒体をローカルのドライブ装置で利用するクライアントコンピュータで当該記憶媒体への更新情報を保持し記憶媒体上に格納する手段と、当該記憶媒体がライブラリ装置に再投入されたとき、前記ライブラリ制御装置が保持する排出された記憶媒体への更新情報とクライアントコンピュータが記録した当該記憶媒体への更新情報を用いて排出されていた間のライブラリ制御装置側での更新を当該記憶媒体に反映する投入媒体同期化手段を設けた。

【0026】上記第四の目的のために、排出された記憶媒体をローカルのドライブ装置で利用するクライアントコンピュータとデータアーカイブサーバ間の通信手段と、その通信手段を用いてクライアントからデータアーカイブサーバでの更新情報を取得する手段と、データアーカイブサーバからクライアントでの更新情報を取得する手段と、更新されたデータを受領する手段を設けた。

【0027】クライアントには、前記可搬媒体の読み出しドライブ装置を備える。

【0028】アーカイブサーバのライブラリ装置は、複数の媒体でバリティグループを構成しそのうちの少なくとも一枚が抜かれても当該媒体へのアクセスを行えるようにする。

【0029】サーバ側でのリード処理はバリティを用いてデータを回復する。

【0030】サーバ側でのライト処理はバリティを更新する。もしくは他の媒体に更新分のみ記録し、更新された領域にダーティフラグを立てておく。当該領域へリード要求が来た場合更新分を記録した媒体を参照する。

【0031】リモート側でのリード処理は、そのままリ

ードする。

【0032】リモート側でのライト処理は、許さない。もしくは、更新用にリザーブした領域に記録し、更新された領域にダーティフラグを立てておく。当該領域へリード要求が来た場合更新分を記録した領域を参照する。

【0033】媒体再挿入後のデータ同期化処理はサーバ側もしくはリモート側どちらか一方のみで更新されたファイルはその更新を有効とし、媒体上で更新を行いダーティフラグを解除する。両方で更新されたファイルは媒体の所有者もしくはシステム管理者の判断もしくは事前設定でサーバ側の更新を有効とするか、リモート側の更新を有効とするか、両方有効とするか、両方無効とするか、二つの新しいファイルとするかを選択する。

【0034】アーカイブサーバは媒体を抜き出した時間および抜き出されたファイルの情報を保持し、当該ファイルへの抜き取り後のアクセスを記録し、クライアントからの要求があった場合その情報をクライアントに送出しファイルの同期化を行う。また、クライアント側でも当該ファイルへのアクセスを記録し、サーバへその情報を送出する手段を備え、当該ファイルの同期化を行う。

【0035】サーバにファイルロック機能を設けリモート側での更新前に通信で更新対象ファイルのロックを確保する。サーバでは抜き取り用媒体作成後の更新を他の媒体に記録し、更新された領域にダーティフラグを立てておく、リモート側でファイルを使用する前に通信でサーバに問い合わせ更新分をもらい媒体にその更新を反映する。リモート側は更新分を通信でサーバに送り、ロックを解除する。

【0036】

【発明の実施の形態】本発明の対象となるデータアーカイブシステムの全体構成を図1に示す。本コンピュータシステムはデータアーカイブサーバ100と、可搬媒体ライブラリ装置200と、クライアントコンピュータ300と、クライアントコンピュータ300に接続し可搬媒体ライブラリ装置200で扱う記憶媒体を利用できるドライブ装置500と、クライアントコンピュータ300とデータアーカイブサーバ100を接続するネットワーク400で構成する。さらにデータアーカイブサーバ100とネットワークで接続されていないリモートコンピュータ600とそれに接続するドライブ装置550も本システムに含めることができる。

【0037】本発明の基本的な特徴は可搬媒体ライブラリ装置200は媒体排出・投入口210を持ち、クライアントコンピュータ300が指定したデータを含む記憶媒体を排出する。排出された記憶媒体をクライアントコンピュータ300のドライブ装置500やリモートコンピュータ600のドライブ装置550を用いて読み書きすることである。

【0038】図2に示すように、データアーカイブサーバ100はクライアントコンピュータ300と接続する

ためのネットワーク制御110と、ライブラリ制御装置200とライブラリ制御のための命令やアーカイブデータなどを送るライブラリ制御130と、クライアントコンピュータ300からくるライブラリ装置にデータ蓄積し、読みだすための要求を解釈し、その実行のためライブラリ制御装置220を生成するデータアーカイブサービス120からなる。

【0039】また、可搬記憶装置ライブラリ200はライブラリ制御装置220とドライブ装置280と媒体投入口210とチェンジャ290と媒体収納庫295で構成する。

【0040】チェンジャ290は媒体収納庫295に格納された記憶媒体を、ライブラリ制御装置220の送出する命令に従い、ドライブ装置280-1~5もしくは媒体投入・排出口210に搬送する。

【0041】図7はクライアントコンピュータ300の構成図である。クライアントコンピュータ300は可搬媒体ライブラリ装置から排出された記憶媒体にアクセスするためのローカルなドライブ装置500を持つ。ネットワークでデータアーカイブサーバ100に接続されていないリモートコンピュータ600も図8に示すようにドライブ装置550を持つ。これらのドライブ装置はリードはできるがライトはできない場合とリードライトともに可能な場合がある。一般にリードのみのドライブ装置のほうが安価であり、かつあとでのべるデータ同期化の問題が発生しない。

【0042】〔記憶媒体のバリティグループと論理ボリューム〕図4、5、6を用いて冗長性を付加した記憶媒体の基本的構成と障害発生時のアクセス方法について説明する。記憶媒体700は固定長のセクタと呼ばれる単位に分割されており、セクタ番号を用いてアクセスする。本可搬媒体ライブラリ装置200では、記録媒体700が障害をおこしてアクセスできなくなった場合でもその障害の影響をクライアントコンピュータ300に与えない高い信頼性を提供する。そのため、図4に示すように複数の記憶媒体700-1、700-2、700-3、700-4を組としてそのうちの一枚700-4にその他の記憶媒体上のデータから計算される冗長なコードを記録しておく。この組をバリティグループ800と呼ぶ。この冗長なコードの典型的な例として、700-1と700-2と700-3のデータすべての排他的論理和で計算されるコードであるバリティが用いられる。このバリティは記憶媒体のセクタ毎に計算する。ライブラリ制御装置220はバリティグループ800を単位としたアクセスをクライアントコンピュータ300に提供する。このアクセス単位を論理ボリュームと呼ぶ。論理ボリュームの中は固定長のセクタに分割され、セクタサイズが同一の場合は記憶媒体700-1~700-3のセクタと1対1の対応がつけられる。この対応はライブラリ制御装置220の構成情報保持手段241が保持す

る記憶媒体700とバリティグループ800の構成情報とセクタ対応規則によって決定される。セクタ対応規則は記憶媒体700-1の先頭セクタから記憶媒体700-3の末尾セクタまで順次対応付けしてもよいし、Pattersonの論文でRAID-4として説明されているようにある一定の単位でローテーションを行う対応付けを行ってもよい。いずれにしても、このバリティグループ800でクライアントコンピュータ300が使用できる容量はこの場合、記憶媒体3枚分となる。

【0043】〔バリティを用いた障害記録媒体へのアクセス処理方法〕バリティグループのなかの一枚の記憶媒体700-2が障害をおこした場合、その記憶媒体上のデータは、図5に示すようにバリティグループの残りの記憶媒体700-1、700-3、700-4の排他的論理和を計算することで再生成する。障害媒体700-2へのライトアクセス処理は図6に示すように対応するバリティ(700-4のP)を700-1のデータD1と400-3のデータD3と400-4の旧バリティ(Pold)と400-2へ書くべきデータ(D2new)の排他的論理和で計算される新バリティ(Pnew)で更新しておく。こうしておけば以降、700-2にリードアクセスがあった場合、図5に示した方法でD2newを再生成することができる。

【0044】この処理はライブラリ制御装置220の障害媒体アクセス手段246で行う。

【0045】以下、本システムでのデータアクセス手順を説明する。

【0046】〔ライブラリに媒体があるばいのアクセス処理〕図9にフローを示す。

【0047】クライアント300がデータアーカイブサーバアクセス制御310のアクセス要求送出手段312を用いてアクセス要求をネットワーク制御340からデータアーカイブサーバ100に送出する(1000)。

【0048】データアーカイブサーバ100のアクセス要求管理手段121が要求を解釈しチェンジャ制御命令を生成(1010)。

【0049】アクセスはリードかライトかをアクセス要求管理手段で判定(1020)。

【0050】リードの場合、チェンジャ290が当該データ媒体をドライブ280に装填(1030)。

【0051】アクセス要求管理手段121がドライブ制御命令を生成しライブラリ制御装置220に送出(1040)。ライブラリ制御装置220の媒体アクセス制御240がドライブ制御260を介して要求データをリードする命令をドライブ装置280に送出し、その結果ドライブ装置280で読み取ったデータをデータアーカイブサーバ100を介してクライアント300に送出(1050)。

【0052】ライトの場合、チェンジャ290が当該データ媒体とそれに対応するバリティ媒体をドライブ28

0に装填(1060)。アクセス要求管理手段121がドライブ制御命令を生成しライブラリ制御装置220に送出(1070)。ライブラリ制御装置220の媒体アクセス制御240がドライブ制御260を介してデータとパリティを更新する命令をドライブ装置280に送出し、その結果ドライブ装置280でデータとパリティを更新し、その後データアーカイブサーバ100を介してクライアント300に終了報告(1080)。

【0053】[ライブラリに媒体がない場合のアクセス処理]クライアント300がアクセス要求をデータアーカイブサーバ100に送出(1100)。データアーカイブサーバ100のアクセス要求管理手段121が要求を解釈しチェンジャ制御命令を生成(1110)。アクセス要求管理手段121がアクセスはリードかライトかを判定(1120)。

【0054】リードの場合、ライブラリ制御装置220の媒体アクセス制御240が構成情報保持手段241、排出・複製媒体アクセス制御手段246を用いて同一パリティグループの媒体を同定し、上記チェンジャ制御命令を変更して、チェンジャ290が当該媒体と同一パリティグループの媒体をドライブ280に装填(1130)。アクセス要求管理手段121がドライブ制御命令を生成(1140)。パリティを用いて図5の方法でデータを再生成(1150)。再生成したデータを、データアーカイブサーバ100を介してクライアント300に送出(1160)。

【0055】ライトの場合、ライブラリ制御装置220の媒体アクセス制御240が構成情報保持手段241、排出・複製媒体アクセス制御手段246を用いて同一パリティグループの媒体を同定し、上記チェンジャ制御命令を変更して、チェンジャ290が当該媒体と同一パリティグループの媒体をドライブ280に装填(1170)。アクセス要求管理手段121がドライブ制御命令を生成(1180)。図6の方法でパリティを更新(1190)。

【0056】データアーカイブサーバ100を介してクライアント300に終了報告を送出(1200)。

【0057】[媒体排出処理]クライアント300がボリューム取り出し要求をデータアーカイブサーバに送出(1201)。

【0058】データアーカイブサーバ100の媒体取り出し・投入要求管理手段123が要求を解釈し媒体排出コマンドをライブラリ制御装置220に送出(1210)。

【0059】媒体アクセス制御240がチェンジャ290に当該媒体を媒体投入・排出口に移動するコマンドを送出し、媒体排出・複製情報保持手段242に記憶(1220)。チェンジャ290が当該媒体を媒体投入・排出口210へ搬送(1230)。クライアント300に終了報告(1240)。

【0060】[媒体複製処理]クライアント300がボリューム複製要求をデータアーカイブサーバに送出(1300)。データアーカイブサーバ100の媒体複製要求管理124が要求を解釈し媒体複製コマンドをライブラリ制御装置220に送出(1310)。媒体アクセス制御240がチェンジャ290に当該データ媒体と複製用媒体をドライブ500に移動するコマンドを送出し、媒体排出・複製情報保持手段242に記憶(1320)。データ媒体から複製用媒体へコピー(1330)。チェンジャ290が複製用媒体を媒体投入・排出口210へ搬送(1340)。クライアント300に終了報告(1350)。

【0061】[クライアントの排出媒体アクセス処理]媒体をドライブ500に装填(1400)。ローカルドライブアクセス制御320ドライブアクセスコマンドを生成(1410)。アクセスはリードかライトかを判定(1420)。

【0062】リードの場合、ドライブ制御330がアクセスコマンドをドライブ500に送出(1430)。ドライブ500から当該データをリード(1440)。

【0063】ライトの場合、同期化制御322で更新情報を登録(1450)。ドライブ制御330がアクセスコマンドをドライブ500に送出(1460)。ドライブ500がデータをライト(1470)。

【0064】[リモート同期化サポート時のクライアントの排出媒体アクセス処理]媒体をドライブ500に装填(1500)。ローカルドライブアクセス制御320ドライブアクセスコマンドを生成(1510)。アクセスはリードかライトかを判定(1520)リードの場合、リモート同期化制御612がデータアーカイブサーバに更新の有無を問い合わせ(1530)。更新あるかどうかを判定(1540)更新がなかった場合、ドライブ制御330がアクセスコマンドをドライブ500に送出(1550)。ドライブ500から当該データをリード(1560)。

【0065】更新があった場合、データアーカイブサーバ100から更新データを取得(1570)。

【0066】ドライブ制御330がアクセスコマンドをドライブ500に送出(1580)。ドライブ500に更新データをライト(1590)。データアーカイブサーバ100で当該セクタの更新情報を削除(1600)。

【0067】ライトの場合、同期化制御322で更新情報を登録(1610)。データアーカイブサーバ100に更新情報を送出(1620)。データアーカイブサーバ100で当該セクタの更新情報を登録(1630)。ドライブ制御330がアクセスコマンドをドライブ500に送出(1640)。ドライブ500がデータをライト(1650)。

【0068】[媒体再投入後の同期化処理]データアー

クライアントコンピュータ 100 に媒体投入を通知 (1700)。媒体をドライブ装置 280 に装填 (1710)。投入媒体同期化制御 249 がリモート側の更新情報とサーバ側の更新情報から同期化処理を実行 (1720)。

【0069】

【発明の効果】本発明によれば、アーカイブデータのネットワーク転送処理なしでアーカイブしたデータを取り出し、持ち運べる集中型データアーカイブサーバシステムが可能となる。

【0070】また、ライブラリ装置を利用したデータアーカイブサーバで記憶媒体が取り外されても当該媒体上のデータへのアクセスを処理することができ、共有データの可用性を高めることができる。

【0071】さらに、上記のシステムで取り出した記憶媒体のデータに対する更新と取り出し中のデータアーカイブサーバ側でのデータ更新を、その記憶媒体がライブラリ記憶装置に再投入された後に同期化することができるので取り出している間にデータアーカイブサーバ側もしくは排出された媒体をローカルに利用しているクライアントコンピュータ側のいずれかでライトアクセスを禁止することなしにデータの整合性を保つことができる。

【0072】排出された媒体をローカルに利用しているクライアントコンピュータとデータアーカイブサーバ間で更新情報を通信することにより、リアルタイムで同期化することができるので比較的低速の通信手段しかない遠隔地でも最新のデータを利用することができ、また、遠隔地で更新されたデータを他のクライアントが即座に共有することができる。

【図面の簡単な説明】

【図 1】本発明の実施例のコンピュータシステムの構成図。

【図 2】本発明の実施例のデータアーカイブサーバの構成図。

【図 3】本発明の可搬媒体ライブラリ装置の構成図。

【図 4】バリティグループの構成図

【図 5】バリティを用いたデータ回復処理。

【図 6】バリティを用いたデータ更新処理。

【図 7】本発明の実施例のクライアントコンピュータの構成。

【図 8】本発明の実施例のリモートコンピュータの構成。

【図 9】ライブラリ装置に媒体がある場合のデータアーカイブサーバでのアクセス処理のフローチャート。

【図 10】ライブラリ装置に媒体がない場合のデータアーカイブサーバでのアクセス処理のフローチャート。

【図 11】ライブラリ装置での媒体排出処理のフローチャート。

【図 12】ライブラリ装置での媒体複製処理のフローチャート。

【図 13】クライアントコンピュータにおけるライブラリ装置から排出された媒体のアクセス処理のフローチャート。

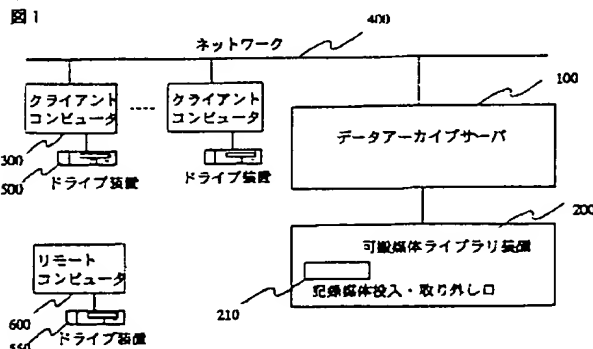
【図 14】遠隔地でリアルタイムにデータの同期化を行う場合のクライアントコンピュータにおけるライブラリ装置から排出された媒体のアクセス処理のフローチャート。

【図 15】ライブラリ装置に排出された媒体を再投入する場合の同期化処理のフローチャート。

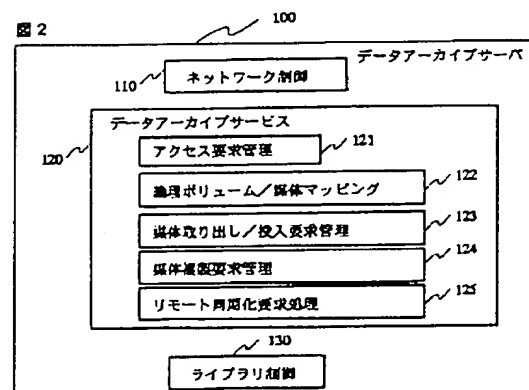
【符号の説明】

100…データアーカイブサーバ、200…可搬媒体ライブラリ装置、300…クライアントコンピュータ、400…ネットワーク、500…クライアントコンピュータのローカルなドライブ装置、600…リモートコンピュータ、700…記憶媒体、800…バリティグループ、210…媒体投入・排出口、220…ライブラリ制御装置、280…可搬媒体ライブラリ装置のドライブ装置、290…媒体チェンジャ、295…媒体収納庫。

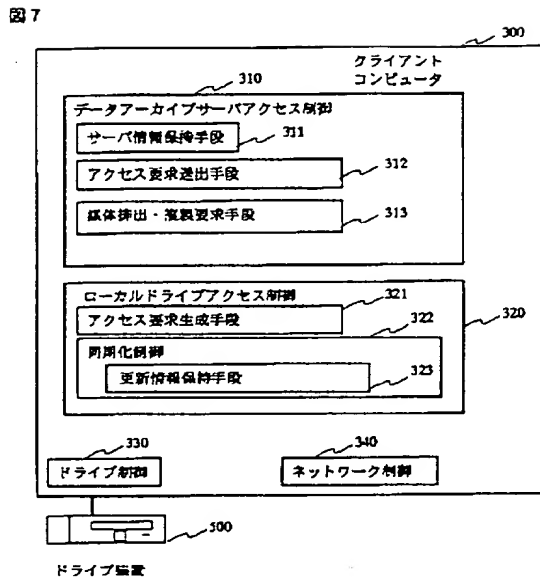
【図 1】



【図 2】

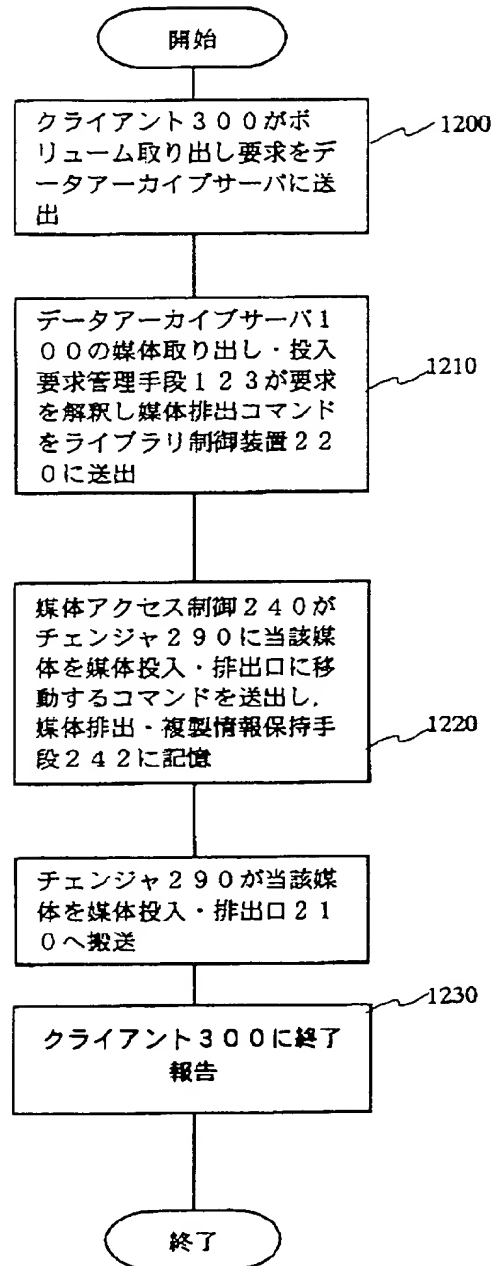


【図7】

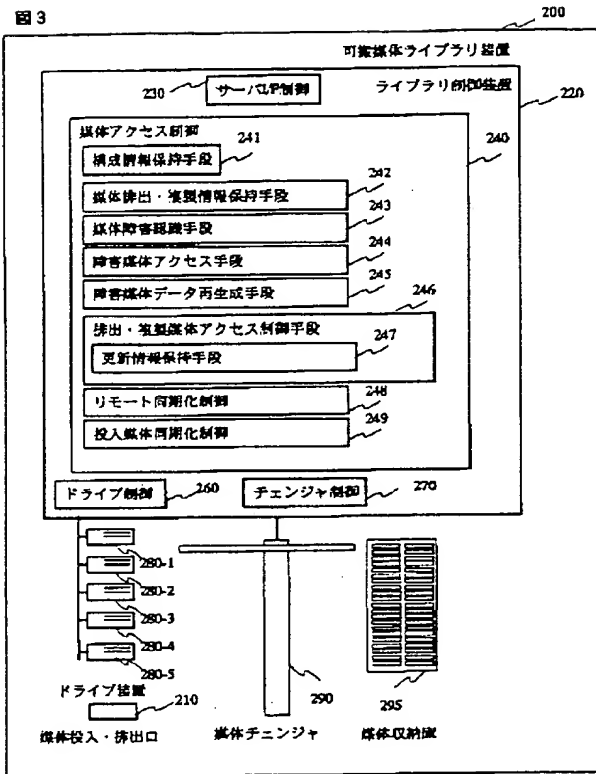


【図11】

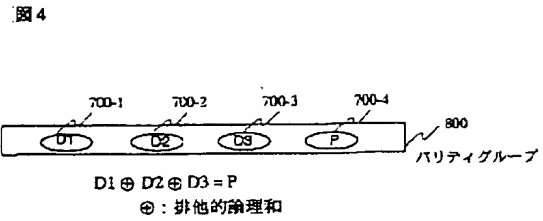
図11



【図 3】

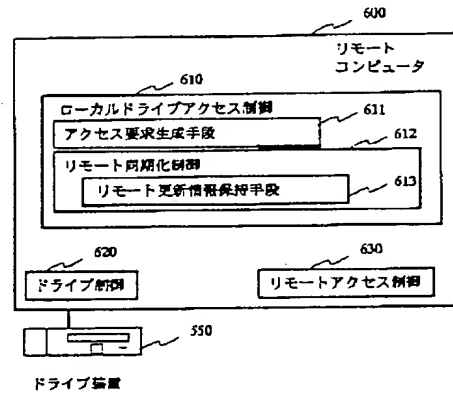


【図 4】

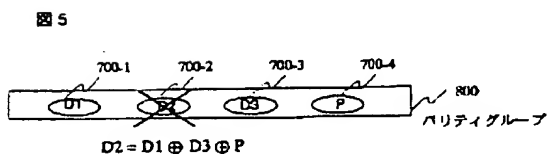


【図 8】

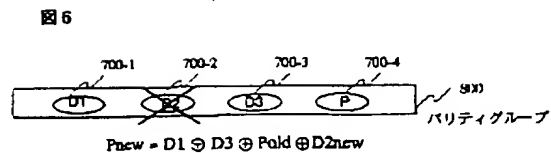
図 8



【図 5】

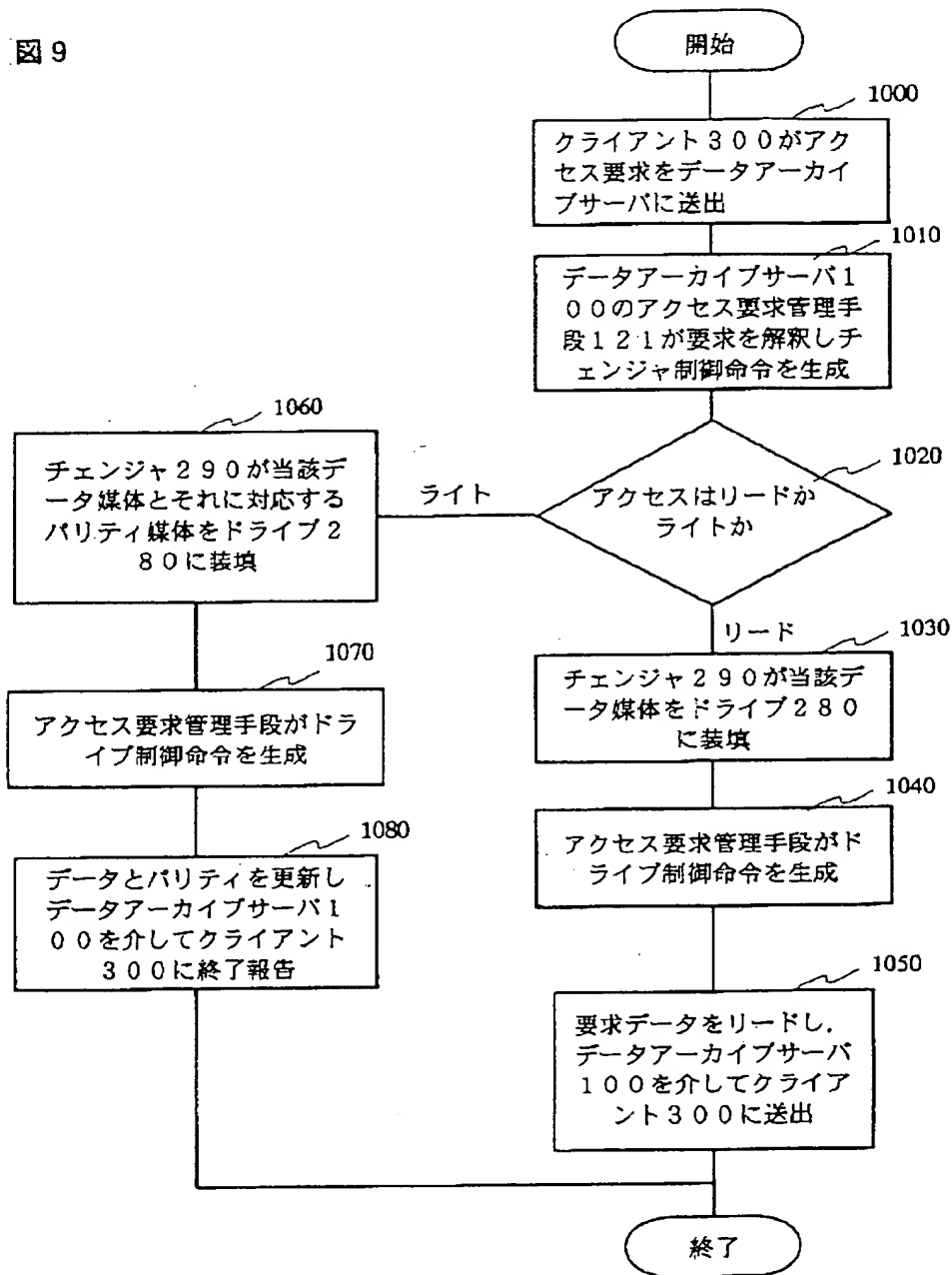


【図 6】



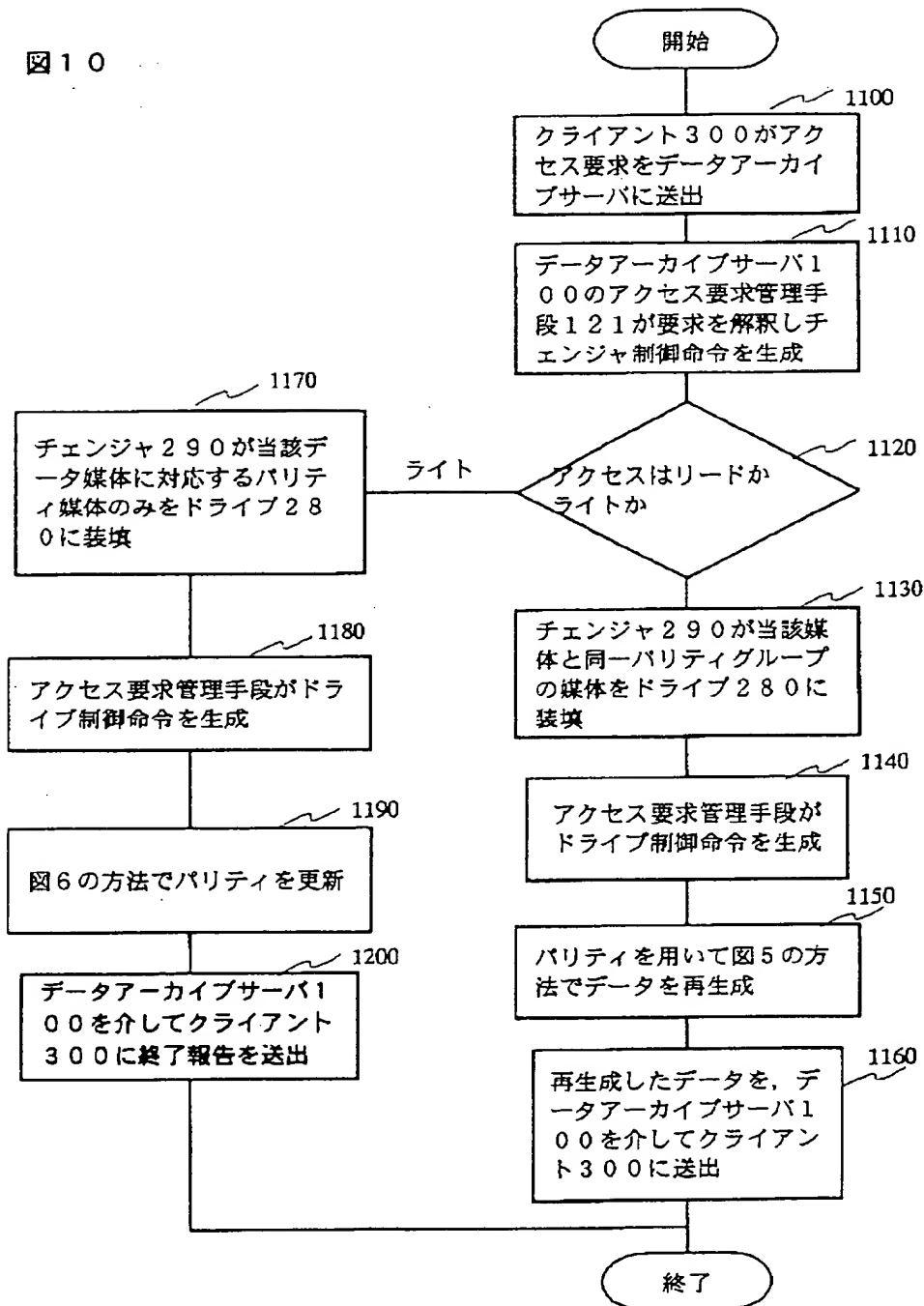
【図9】

図9



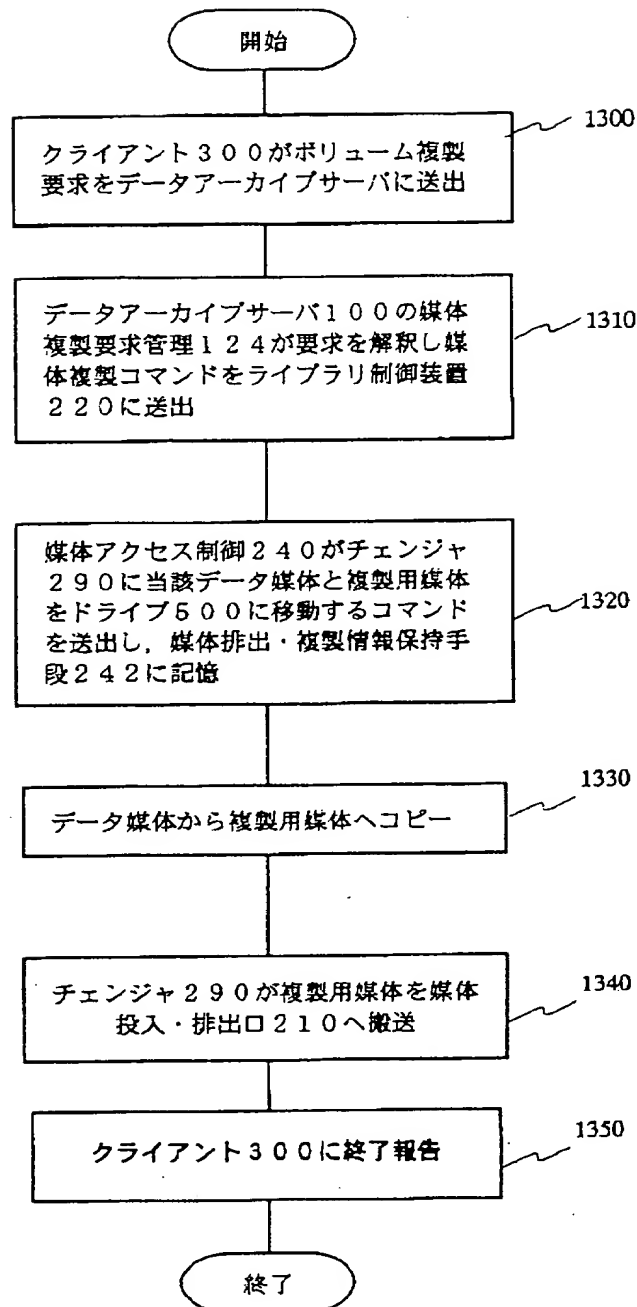
【図10】

図10



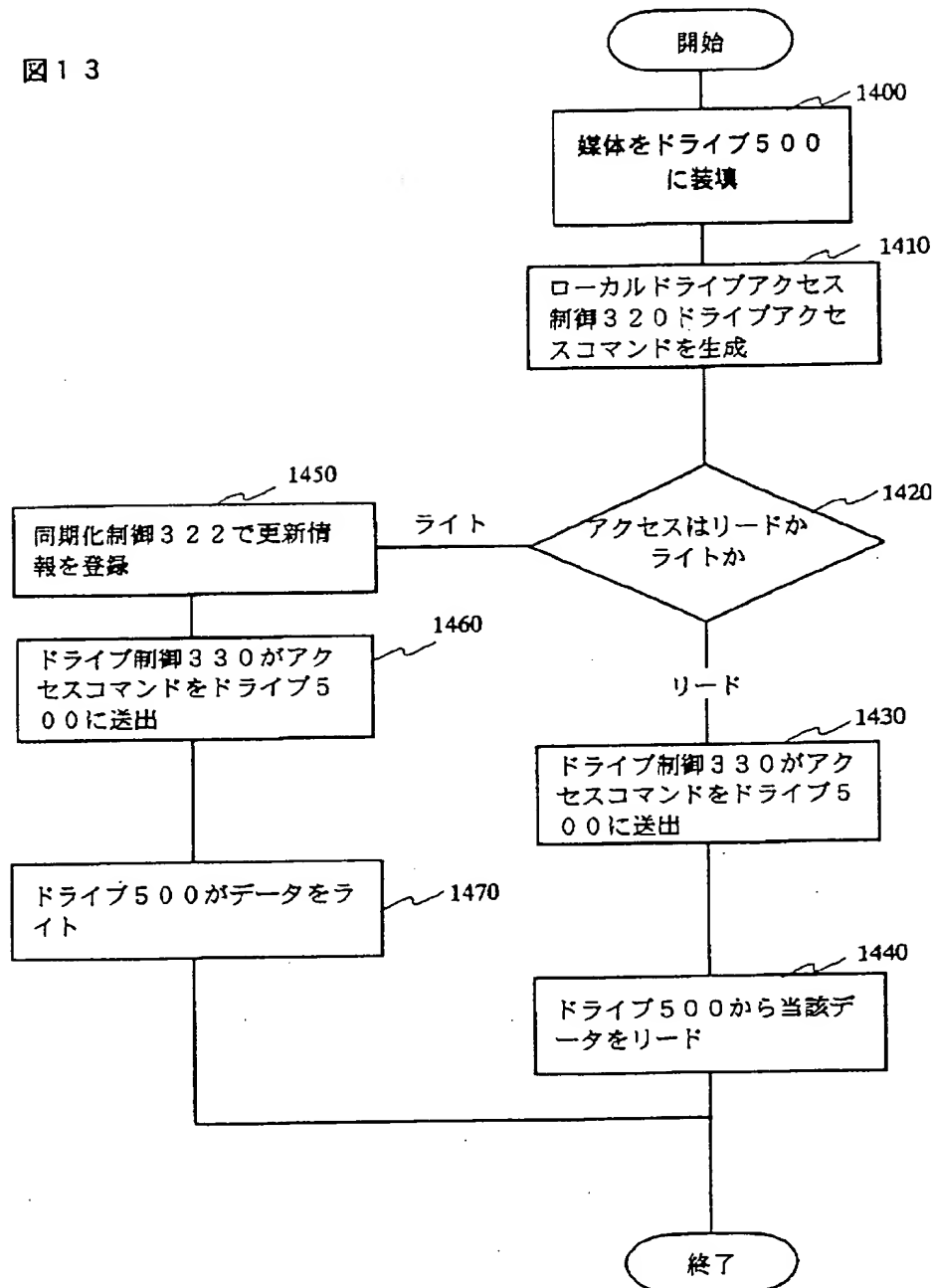
【図12】

図12

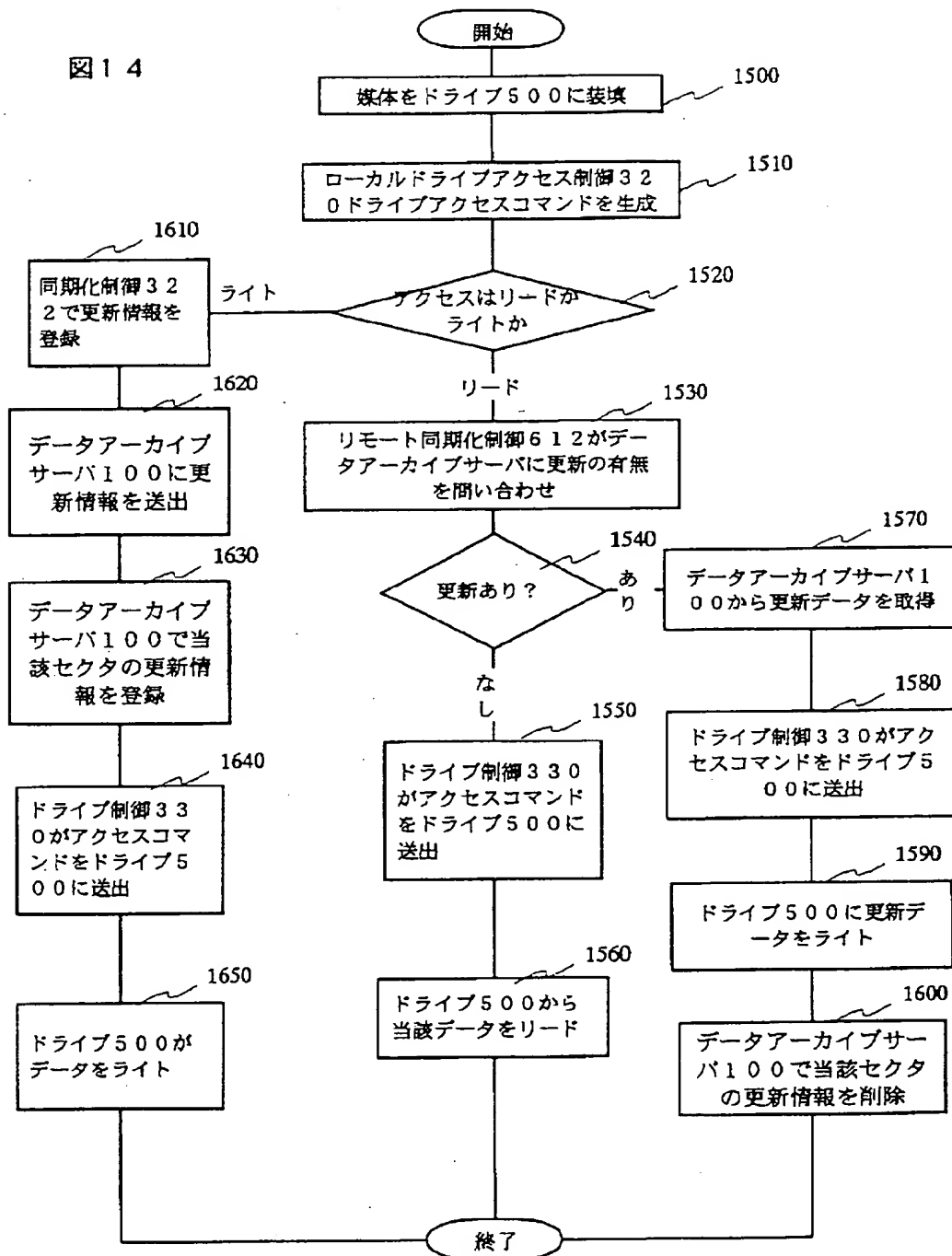


【図13】

図13

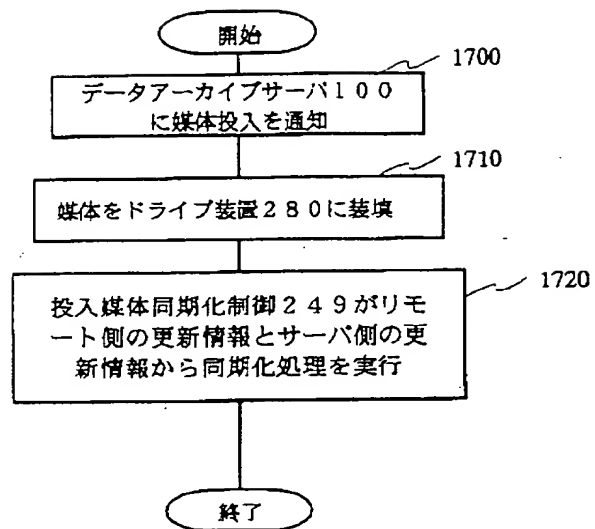


【図14】



【図 15】

図 15



フロントページの続き

(72)発明者 山本 彰
神奈川県川崎市麻生区王禅寺1099番地 株
式会社日立製作所システム開発研究所内